#15

PATENT

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In Re Application of

Appellant	:	Howard T. Olnowich
Serial No.	:	09/394,564
Filed	:	10 Sep 1999
Examiner	:	P. Bataille
Group	:	2186
Entitled	:	Memory Controller for Controlling Memory Accesses Across Networks in Distributed Shared Memory Processing Systems (as amended)
Docket No.	:	EN997080B

Assistant Commissioner for Patents
Washington, D.C. 20231

ATTENTION: Board of Patent Appeals and Interferences

APPELLANT'S BRIEF (37 C.F.R. § 1.192)

This brief is in furtherance of the Notice of Appeal, filed in this case on 2 April 2002.

CERTIFICATE OF MAILING (37 C.F.R. § 1.8(A))

I hereby certify that this correspondence is, on the date shown below, being deposited with the United States Postal Service with sufficient postage as first class mail in an envelope addressed to the Commissioner of Patents and Trademarks, Washington, D.C. 20231.

Name: JUDITH A. BECKSTRAND
Date: 28 MAY 2002 Signature: Judith A. Beckstrand

The fees required under 37 C.F.R. §1.17, and any required petition for extension of time for filing this brief and fees therefor, are dealt with in the accompanying TRANSMITTAL OF APPEAL BRIEF.

This brief is transmitted in triplicate.

This brief contains these items under the following headings, and in the order set forth below:

I REAL PARTY INTEREST

II RELATED APPEALS AND INTERFERENCES

III STATUS OF CLAIMS

IV STATUS OF AMENDMENTS

V SUMMARY OF INVENTION

VI ISSUES

VII GROUPING OF CLAIMS

VIII ARGUMENTS

- ARGUMENT: VIIID REJECTIONS UNDER 35 U.S.C. 103

IX APPENDIX OF CLAIMS INVOLVED IN THE APPEAL

The final page of this brief bears the practitioner's

signature.

I REAL PARTY INTEREST

The real party in interest in this appeal is International Business Machines Corporation, Armonk, New York.

II RELATED APPEALS AND INTERFERENCES

No appeals or interferences will directly affect, or be directly affected by, or have a bearing on the Board's decision in this appeal.

III STATUS OF CLAIMS

A. TOTAL NUMBER OF CLAIMS IN APPLICATION

Claims in the application are: 1 and 31-39.

B. STATUS OF ALL THE CLAIMS

1. Claims canceled: None

2. Claims withdrawn from consideration but not canceled: None

3. Claims pending: 1 and 31-39

4. Claims allowed: None

5. Claims rejected: 1 and 31-39

C. CLAIMS ON APPEAL

The claims on appeal are: 1 and 31-39

IV STATUS OF AMENDMENTS

The status of any amendment filed subsequent to the final rejection is, insofar as understood by appellant, as follows:

No amendments to the specification or claims have been

filed subsequent to the final rejection, dated 29 Nov 2001. Applicant filed 22 Jan 2002 a response which has not been entered, but included no amendments to the claims or specification.

V SUMMARY OF INVENTION

Referring primarily to Figures 2A and 2B of the present application, a cache coherency system for a shared memory parallel processing system including a plurality of processing nodes 30, 34 (page 21, lines 1-11, or page 21:1-11), comprising:

a multi-stage communication network 20 (pages 23:15-16; 25:4-10; 32:11-21) for interconnecting said processing nodes 30, 34;

each said processing node 30, 34 including a unique section 222 (pages 22:9; 23:23-24:13; 34:25-35:3; 41:7-42:8)of shared memory 54 which is not a cache 100, 204;

each said processing node 30, 34 including one or more caches 100, 204 (page 16:17-18; 58:28-32) for storing a

plurality of cache lines (page 59:3,28-22);

a cache coherency directory 32 (pages 36:4-15; 16:18-19) which is distributed to each of said nodes 30, 34 for tracking which of said nodes have copies of each cache line (pages 45:11-19; 56:12; 63:26-64:3); and

an adapter 10 for storing (line 216) changed data immediately (write-through mode: page 72:4-9) to said unique section 222 of shared memory 54 regardless of which of said nodes 30, 34 is changing the data and which of said nodes includes the section 222 of shared memory 54 to be changed, such that said shared memory 54 always contains the most recent data (page 72:10-73:32; 82:6-10; 83:12-84:17).

VI ISSUES

Whether claims 1 and 31-390 are unpatentable under 35 U.S.C. 103 over Gupta et al. (U.S. Patent 5,535,116) in view of Hagersten et al. (U. S. Patent 5,887,138).

VII GROUPING OF CLAIMS

Claims 1 and 31-39, stand or fall together.

VIII ARGUMENTS

[...contentions, citations of authority, statutes, and parts of the record relied upon...]

ARGUMENT: VIIID REJECTIONS UNDER 35 U.S.C. 103

Claims 1 and 31-39 have been rejected under 35 U.S.C. 103(a) over Gupta et al. (U.S. Patent 5,535,116, hereafter Gupta) in view of Hagersten et al. (U.S. Patent 5,887,138, hereafter Hagersten).

Appellant contends that the Examiner has not established a prima facie case of obviousness. Such is established when the Examiner provides (1) one or more references (2) that were available to the inventor and (3) that teach (4) a suggestion to combine or modify the

references, (5) the combination or modification of which would appear to be sufficient to have made the claimed invention obvious to one of ordinary skill in the art.

Appellant argues that the Examiner has not satisfied requirement 5, for the combination which he suggests does not teach the claimed invention. The decision in *In re Ratti*, 270 F.2d 810, 123 USPQ 349 (C.C.P.A. 1959) indicates that an obviousness rejection is not appropriate if substantial reconstruction or redesign of the prior art references is necessary to arrive at the invention. As will be demonstrated hereafter, in order for Hagersten and Gupta to be combined to suggest appellant's claimed invention would require substantial reconstruction or redesign of the references. Further, the art references of Gupta and Hagersten do not meet the substantially same function, way, and result test of *Hilton Davis Chemical Co. v. Warner-Jenkinson Co.* 62 F.3d 1512, 35 USPQ 2d 1641, 1645 (Fed. Cir. 1995) and *Pannwalt Corp. v. Durand-Wayland, Inc.*, 833 F.2d 931, 4 USPQ2d 1737 (Fed. Cir. 1987)(en bac), *cert. denied*, 485 u.s. 961 (1988).

Each of appellant's claims require that the most recent data be stored immediately (that is, write-through) to a

unique section of shared memory which is not a cache and which is distributed to a plurality of nodes communicating over a multi-stage network. Neither Hagersten nor Gupta teach such (that is, do not perform substantially the same function in the same way to achieve the same result) and substantial reconstruction or redesign of these references would be required which would only be possible through hindsight reasoning based on appellant's own teachings.

With respect to Gupta, as the Examiner states in the Office Action, "Gupta fails to specifically teach the features... each processing node including a unique section of shared memory which is not a cache and an adapter for storing changed data to each unique section so that each section contains the most recent data."

However, the Examiner states that Hagersten teaches these features, and appellant traverses.

The Examiner correctly states that Hagersten teaches a plurality of processing nodes interconnected through an interconnecting network; each processing node including multiple processors, multiple cache memories and local memory, all local memories or memory portions of the processing nodes collectively form a distributive shared

memory which may be accessed in non-uniform memory architecture (NUMA) fashion, i.e. each of said multi-processing nodes includes an addressable portion or local memory modules of the global system memory or sub-divided portion of the global physical system memory, and each processing node having said local memory is capable of storing valid and shared copies of requested ones of data signals stored in the main memory modules. The Examiner further correctly states that Hagersten in many instances indicates that each node of the plurality of nodes shares a shared distributed memory wherein sections or addressable locations of the shared distributed memory are accessible to more than one of the processing nodes. Hagersten's system does not simply provide bus operations, but provides a coherency system maintaining coherency between the memories within the multi-processing nodes communicating via an interconnecting network.

Further, the Examiner is correct regarding Hagersten having "external device owning exclusive copy can unilaterally modify the copy without having to inform other entities". However, in Hagersten, this is for an inter-node or processor-bus transaction within a subnode since all processors within a node have respective cache memory, and common local memory or system memory portion.

As will be described hereafter, Hagersten's system for maintaining internode coherency is not at all similar to that taught and claimed appellant.

1) Hagersten Does Not Teach the Cache Coherency Aspect of Appellant's Claims

Hagersten teaches a write-back coherency operation which floods the network with excessive traffic, whereas the present application teaches a write-thru cache, which provides a better solution and requires much less network traffic.

For the present invention, the write-thru feature is claimed in every independent claim with statements similar to: "changing said shared memory, wherein changed data is stored immediately to shared memory regardless of which of said nodes is changing the data and which of said nodes includes the section of shared memory to be changed, wherein said shared memory always contains the most recent data."

On the other hand, Hagersten teaches write-back coherency, where a "requesting node" must query a "home node" for the location of which node and cache contains the

most recent data. The "home node" searches for the location of the cache holding the requested data amongst "slave nodes" (i.e., all nodes having a cached copy of the requested data). Then, all slaves nodes must respond to the requesting node, where the "owning" slave node sends the latest value of the requested data and the other slave nodes acknowledge that they also have a copy of the data. Hagersten calls this coherency process a "three-hop communication" [col.5, line 42]. These hops are as follows: hop 1) requesting node to home node, hop 2) home node to all slave nodes having a copy of the data in search of the most recent copy, and hop 3) all slaves nodes to the requesting node. Hagersten explains this three-hop process in col.5. lines 29-43.

The present invention does not require a three-hop or four-hop method (four hop is referenced by Hagersten at col. 4, line 45), where the cache of any slave node can store the most recent data. In contrast, the present application teaches a write-thru cache coherency method, whereby any node changing the data immediately writes (thus, writes through) the most recent data back to the shared memory at the home node, as is set forth, for example, in claim 1 at lines 14-19. The result is that the shared memory at the home node (to use Hagersten's word) always has the most

recent data, and the home node never has to search any other nodes or caches for where the most recent data is stored. Because of this feature of appellant's invention, and in contrast with Hagersten, the present invention is capable of implementing a two-hop method as follows: hop 1) a requesting node requests the most recent data of the home node, and hop 2) the home immediately returns the most recent data from its shared memory to the requesting node. This provides the quickest response possible for fetching global data.

2) Hagersten Does Not Teach a Cache Coherency Network

Hagersten teaches point-to-point network 14 as shown in Figure 1, where each node 12A-D communicates over a direct link to each other node 12A-D over network 14. Hagersten network 14 is designed to allow his cache coherency method to work more efficiently. Suppose, for example, Hagersten node 12A is the requesting node that requests data, node 12B is the home node, and nodes 12C-D are the slaves nodes. For hop 1, a network transaction takes place between 12A to 12B on network 14. For hop 2, a network transaction takes place between 12B to 12C and 12D on network 14; however, since 12B has parallel point-to-point connections through network 14

to 12C and 12D, the communications can take place simultaneously while leaving the rest of the network 14 connections open for other network traffic. For hop 3, network transactions take place between 12C to 12A and 12D to 12A on network 14; however, since 12A has parallel point-to-point connections through network 14 to 12C and 12D, the communications can take place simultaneously while leaving the rest of the network 14 connections open for other network traffic. Thus, Hagersten's cache coherency method works faster on his point-to-point network since the network is designed for parallel communications over direct links between nodes.

However, the drawback of Hagersten's network is that it is not a good network for expanding to many nodes, because as the number of nodes n increases, the number of point-to-point connecting links (and therefore cables) increases by n[(n-1)!] and the complexity of the System Interface block 24 of Figure 1 increases by n. This means Hagersten's method becomes very unwieldy with an increase in nodes.

In contrast, the present invention is designed for a different type of network, a multi-stage network shown in Figure 4 of the present invention, where each node has one and only one connection to the network and the network is

designed for expandability to hundreds or thousands of nodes. Note that each claim of the present clearly states that the present invention is for a multi-stage network, which is not a point-to-point network. Thus, Hagersten's solution is a limited solution that applies well to a small number of nodes and works more efficiently on a parallel point-to-point network.

Hagersten observes that "in alternative embodiments other networks may be used" (Col. 9, lines 7-8). If Hagersten's method is run on a multi-stage network, all communications of home nodes to slave nodes (hop 2) and slave nodes to requesting nodes (hop 3) would have to take place in a serial fashion as each node has one and only one serial connection to a multi-stage network. The serial communication aspect of transfers from or to any given node of a multi-stage network plus the larger number of communications required in searching for the slave nodes and having all slave nodes respond, would cause Hagersten's method to perform very poorly on a multi-stage network.

Appellant's claims require, and neither Hagersten nor Gupta teach that, in a system including a plurality of nodes interconnected over a multistage communication network, changed data be stored immediately at each node to that

unique section of shared memory at each node which is not a cache regardless of which node changes the data and regardless of which node includes the section of shared memory to be changed.

Thus, Hagersten performs in a substantially different way under the function-way-result test using a different structure, and should not be considered an equivalent of the claimed invention to support an obviousness rejection. See *Engel Industries v. The Lockformer Co.*, 96 F.3d 1398, 40 USPQ 2d 1161 (Fed. Cir. 1996).

The claims of the present invention clearly differentiate unique aspects over Hagersten and/or Gupta, those aspects including write-thru caches storing the most recent data directly to a shared memory which is distributed to a plurality of nodes communicating over a multi-stage network.

In the Advisory Action mailed 03/20/2002, paper 13, the Examiner states:

"The... request for reconsideration has been considered but does NOT place the application in condition for

allowance because...: despite the noted distinction between applicant's claimed invention and the cited reference by Hagersten (US 5,887,138), Hagersten covers the features as required in the claims. Specifically Hagersten teaches the required feature, each of a plurality of processing node including a unique section of shared memory which is not a cache (a sub-divided portion or an addressable portion or local memories of the global memory assigned to each processing node (see Columns 8 and 9).

In this characterization of Hagersten, the Examiner is silent as to whether Hagersten teaches, as set forth, for example, in appellant's claim 1: "...an adapter for storing changed data immediately to said unique section of shared memory regardless of which of said nodes is changing the data and which of said nodes includes the section of shared memory to be changed, such that said shared memory always contains the most recent data." This is the write-through aspect of appellant's claims, and as is explained above is not taught by Hagersten or Gupta.

CONCLUSION

Appellant requests that the rejection of claims 1 and 31-39 be reversed.

Shelley M Beckstrand
Attorney for Appellant
Reg. No. 24,886

314 Main Street
Owego, NY 13827

Phone: (607) 687-9913
Fax: (607) 687-7848

IX APPENDIX OF CLAIMS INVOLVED IN THE APPEAL

1. [Twice Amended] A cache coherency system for a shared memory parallel processing system including a plurality of processing nodes, comprising:

a multi-stage communication network for interconnecting said processing nodes;

each said processing node including a unique section of shared memory which is not a cache;

each said processing node including one or more caches for storing a plurality of cache lines;

a cache coherency directory which is distributed to each of said nodes for tracking which of said nodes have copies of each cache line; and

an adapter for storing changed data immediately to said unique section of shared memory regardless of which of said nodes is changing the data and which of said nodes includes the section of shared memory to be changed,

such that said shared memory always contains the most recent data.

31. [Amended] A method for operating a shared memory parallel processing system as a cache coherency system including a plurality of processing nodes, each said processing node including a unique section of shared memory which is not a cache, comprising the steps of:

interconnecting said processing nodes through a multi-stage communication network;

storing at each said processing node a plurality of cache lines in one or more caches;

distributing to each of said processing nodes a cache coherency directory;

tracking in said cache coherency directory which of said processing nodes have copies of each cache line; and

changing said shared memory, wherein changed data is

stored immediately to said unique section of shared memory regardless of which of said nodes is changing the data and which of said nodes includes the section of shared memory to be changed, wherein said shared memory always contains the most recent data.

32. [Amended] A program storage device readable by a machine, tangibly embodying a program of instructions executable by a machine to perform method steps for operating a shared memory parallel processing system including a plurality of processing nodes, each said processing node including a unique section of shared memory which is not a cache, said method steps comprising:

interconnecting said processing nodes through a multi-stage communication network;

storing at each said processing node a plurality of cache lines in one or more caches;

tracking in a cache coherency directory which is distributed to each of said processing nodes which of

said processing nodes have copies of each cache line; and

changing said unique section of shared memory, wherein changed data is stored immediately to shared memory regardless of which of said nodes is changing the data and which of said nodes includes the section of shared memory to be changed, wherein said shared memory always contains the most recent data.

33. [Amended] An article of manufacture comprising:

a computer useable medium having computer readable program code means embodied therein for operating a shared memory parallel processing system including a plurality of processing nodes, each said processing node including a unique section of shared memory which is not a cache, the computer readable program means in said article of manufacture comprising:

computer readable program code means for causing a computer to effect interconnecting said processing nodes through a multi-stage communication network;

computer readable program code means for causing a computer to effect storing at each said processing node a plurality of cache lines in one or more caches;

computer readable program code means for causing a computer to effect tracking in a cache coherency directory which is distributed to each of said processing nodes which of said processing nodes have copies of each cache line; and

computer readable program code means for storing changed data immediately to said unique section of shared memory regardless of which of said nodes is changing the data and which of said nodes includes the section of shared memory to be changed such that said shared memory always contains the most recent data.

34. [Amended] A computer program product or computer program element for operating a shared memory parallel processing system including a plurality of processing nodes, each said node including a unique section of shared memory which is not a cache, according to the steps of:

interconnecting said processing nodes through a multi-

stage communication network;

storing at each said processing node a plurality of cache lines in one or more caches;

distributing to each of said processing nodes a cache coherency directory;

tracking in said cache coherency directory which of said processing nodes have copies of each cache line; and

storing changed data immediately to said unique section of shared memory regardless of which of said nodes is changing the data and which of said nodes includes the section of shared memory to be changed such that said shared memory always contains the most recent data.

35. The cache coherency system of claim 1, further comprising:

a shared memory including a first memory portion for storing unchangeable data and a second memory portion for storing changeable data; and

said cache coherency directory listing which nodes of said plurality of processing nodes have accessed copies of said cache lines in said second memory portion.

36. The cache coherency system of claim 35, each of said plurality of processing nodes being operable for reading, storing, and invalidating said shared memory at any other of said processing nodes.

37. [Amended] The cache coherency system of claim 36, further comprising at a first node of said plurality of processing nodes a memory controller selectively operable first responsive to a request for access to a memory word by first accessing the cache at said first node and, if said requested memory word is not available in said cache, selectively operable second for accessing said memory word selectively from said shared memory regardless of which of said nodes includes the section of shared memory being accessed, and storing said cache line including said memory word to said cache at said first node.

38. [Amended] The cache coherency system of claim 37, said

memory controller further being selectively operable for deleting a cache line from said cache at said first node when said cache is full to provide space for a new cache line to be stored to said cache, and for sending the address of the deleted cache line to an invalidation directory to indicate said node no longer has a copy of said cache line.

39. [Amended] The cache coherency system of claim 37, said memory controller further being selectively operable for sending cache update messages to update corresponding cache lines at all remote nodes having copies of a changed cache line and for receiving cache lines of data from remote nodes for updating the cache at said first node.

2186

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re application of:

Applicant	:	H. T. Olnowich
Serial No.	:	09/394,564
Filed	:	10 Sep 1999
Group No.	:	2186
Examiner	:	P. M. Bataille
Docket	:	EN997080B
For	:	Memory Controller for Controlling Memory Accesses Across Networks In Distributed Shared Memory Processing Systems

Assistant Commissioner For Patents
Washington, D.C. 20231

CERTIFICATE OF MAILING UNDER 37 CFR 1.8(a)

I hereby certify that the following attached correspondence comprising:

Acknowledgment Postcard
Certificate of Mailing
Transmittal of Appeal Brief
Appellant's Brief

is being deposited with the United States Postal Service as first class mail in an envelope addressed to:

Commissioner of Patents and Trademarks
Washington, D. C. 20231

on 28 May 2002 .
(date)

JUDITH A. BECKSTRAND
(Name of person mailing paper or fee)

Judith A. Beckstrand
(Signature of person mailing paper or fee)

PATENT

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In Re Application of

Appellant	:	Howard T. Olnowich
Serial No.	:	09/394,564
Filed	:	10 Sep 1999
Examiner	:	P. Bataille
Group	:	2186
Entitled	:	Memory Controller for Controlling Memory Accesses Across Networks in Distributed Shared Memory Processing Systems (as amended)
Docket No.	:	EN997080B

Assistant Commissioner for Patents
Washington, D.C. 20231

TRANSMITTAL OF APPEAL BRIEF
(PATENT APPLICATION — 37 C.F.R. § 1.192)

1. TRANSMITTAL

Transmitted herewith, in triplicate, is the APPEAL BRIEF in this application, with respect to the Notice of Appeal filed on 2 Apr 2002.

2. STATUS OF APPLICANT

This application is on behalf of other than a small entity.

CERTIFICATE OF MAILING (37 C.F.R. § 1.8(A))

I hereby certify that this correspondence is, on the date shown below, being deposited with the United States Postal Service with sufficient postage as first class mail in an envelope addressed to the Commissioner of Patents and Trademarks, Washington, D.C. 20231.

Name: JUDITH A. BECKSTRAND
Date: 28 may 2002 Signature: Judith A. Beckstrand

---

3. FEE FOR FILING APPEAL BRIEF

Pursuant to 37 C.F.R. § 1.17(c), the fee for filing the Appeal Brief is $320.00 for other than a small entity.

4. EXTENSION OF TERM

The proceedings herein are for a patent application and the provisions of 37 C.F.R. § 1.136 apply. Applicant believes that no extension of term is required. However, if an extension of term is required, please consider this a petition therefor.

5. TOTAL FEE DUE:

The total fee due is:

Appeal brief fee	$ 320.00
Extension fee (if any)	$ ____
Total Fee Due:	$ 320.00

6. FEE PAYMENT:

Charge IBM Deposit Account No. 09-0457 the sum of $ 320.00. A duplicate of this transmittal is attached.

7. FEE DEFICIENCY

This is a request to charge IBM Deposit Account No. 09-0457 for any required additional extension and/or fee, or for any required additional fee for claims.

Shelley M Beckstrand
Attorney for Applicant(s)
Reg. No. 24,886

314 Main Street
Owego, NY 13827

Phone: (607) 687-9913
Fax: (607) 687-7848
Cell phone: (607) 427-0802